



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 02:53 am GMT

PDB ID : 3MX4  
Title : DNA binding and cleavage by the GIY-YIG endonuclease R.Eco29KI inactive variant E142Q  
Authors : Mak, A.N.S.; Lambert, A.R.; Stoddard, B.L.  
Deposited on : 2010-05-06  
Resolution : 2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

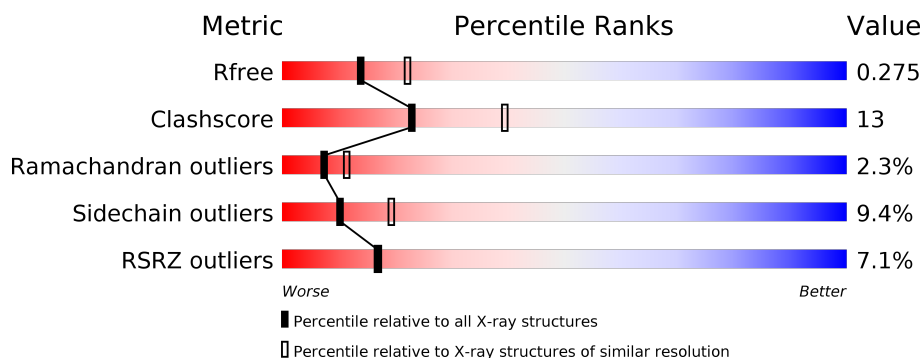
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	235	<div> <div>6%</div> <div> <div></div> <div>69%</div> <div>18%</div> <div>• • 9%</div> </div> </div>
1	B	235	<div> <div>6%</div> <div> <div></div> <div>65%</div> <div>23%</div> <div>• 9%</div> </div> </div>
1	C	235	<div> <div>8%</div> <div> <div></div> <div>62%</div> <div>23%</div> <div>• 11%</div> </div> </div>
1	D	235	<div> <div>6%</div> <div> <div></div> <div>60%</div> <div>23%</div> <div>6% 11%</div> </div> </div>
1	E	235	<div> <div>8%</div> <div> <div></div> <div>66%</div> <div>19%</div> <div>5% 10%</div> </div> </div>
1	F	235	<div> <div>5%</div> <div> <div></div> <div>63%</div> <div>20%</div> <div>6% 11%</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	235	
1	H	235	
2	I	22	
2	K	22	
2	M	22	
2	O	22	
3	J	22	
3	L	22	
3	N	22	
3	P	22	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 17406 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Eco29kIR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	0	0	0
			1677	1069	298	306	4			
1	B	213	Total	C	N	O	S	0	0	0
			1687	1077	299	307	4			
1	C	208	Total	C	N	O	S	0	0	0
			1649	1051	292	302	4			
1	D	208	Total	C	N	O	S	0	0	0
			1634	1041	290	299	4			
1	E	211	Total	C	N	O	S	0	0	0
			1663	1062	297	300	4			
1	F	210	Total	C	N	O	S	0	0	0
			1674	1066	298	306	4			
1	G	208	Total	C	N	O	S	0	0	0
			1663	1062	294	303	4			
1	H	209	Total	C	N	O	S	0	0	0
			1658	1056	295	303	4			

There are 192 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	EXPRESSION TAG	UNP Q46944
A	-19	ALA	-	EXPRESSION TAG	UNP Q46944
A	-18	SER	-	EXPRESSION TAG	UNP Q46944
A	-17	SER	-	EXPRESSION TAG	UNP Q46944
A	-16	HIS	-	EXPRESSION TAG	UNP Q46944
A	-15	HIS	-	EXPRESSION TAG	UNP Q46944
A	-14	HIS	-	EXPRESSION TAG	UNP Q46944
A	-13	HIS	-	EXPRESSION TAG	UNP Q46944
A	-12	HIS	-	EXPRESSION TAG	UNP Q46944
A	-11	HIS	-	EXPRESSION TAG	UNP Q46944
A	-10	SER	-	EXPRESSION TAG	UNP Q46944
A	-9	SER	-	EXPRESSION TAG	UNP Q46944
A	-8	GLY	-	EXPRESSION TAG	UNP Q46944

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	LEU	-	EXPRESSION TAG	UNP Q46944
A	-6	VAL	-	EXPRESSION TAG	UNP Q46944
A	-5	PRO	-	EXPRESSION TAG	UNP Q46944
A	-4	ARG	-	EXPRESSION TAG	UNP Q46944
A	-3	GLY	-	EXPRESSION TAG	UNP Q46944
A	-2	SER	-	EXPRESSION TAG	UNP Q46944
A	-1	SER	-	EXPRESSION TAG	UNP Q46944
A	0	MET	-	EXPRESSION TAG	UNP Q46944
A	1	GLY	-	EXPRESSION TAG	UNP Q46944
A	69	LYS	LEU	ENGINEERED MUTATION	UNP Q46944
A	142	GLN	GLU	ENGINEERED MUTATION	UNP Q46944
B	-20	MET	-	EXPRESSION TAG	UNP Q46944
B	-19	ALA	-	EXPRESSION TAG	UNP Q46944
B	-18	SER	-	EXPRESSION TAG	UNP Q46944
B	-17	SER	-	EXPRESSION TAG	UNP Q46944
B	-16	HIS	-	EXPRESSION TAG	UNP Q46944
B	-15	HIS	-	EXPRESSION TAG	UNP Q46944
B	-14	HIS	-	EXPRESSION TAG	UNP Q46944
B	-13	HIS	-	EXPRESSION TAG	UNP Q46944
B	-12	HIS	-	EXPRESSION TAG	UNP Q46944
B	-11	HIS	-	EXPRESSION TAG	UNP Q46944
B	-10	SER	-	EXPRESSION TAG	UNP Q46944
B	-9	SER	-	EXPRESSION TAG	UNP Q46944
B	-8	GLY	-	EXPRESSION TAG	UNP Q46944
B	-7	LEU	-	EXPRESSION TAG	UNP Q46944
B	-6	VAL	-	EXPRESSION TAG	UNP Q46944
B	-5	PRO	-	EXPRESSION TAG	UNP Q46944
B	-4	ARG	-	EXPRESSION TAG	UNP Q46944
B	-3	GLY	-	EXPRESSION TAG	UNP Q46944
B	-2	SER	-	EXPRESSION TAG	UNP Q46944
B	-1	SER	-	EXPRESSION TAG	UNP Q46944
B	0	MET	-	EXPRESSION TAG	UNP Q46944
B	1	GLY	-	EXPRESSION TAG	UNP Q46944
B	69	LYS	LEU	ENGINEERED MUTATION	UNP Q46944
B	142	GLN	GLU	ENGINEERED MUTATION	UNP Q46944
C	-20	MET	-	EXPRESSION TAG	UNP Q46944
C	-19	ALA	-	EXPRESSION TAG	UNP Q46944
C	-18	SER	-	EXPRESSION TAG	UNP Q46944
C	-17	SER	-	EXPRESSION TAG	UNP Q46944
C	-16	HIS	-	EXPRESSION TAG	UNP Q46944
C	-15	HIS	-	EXPRESSION TAG	UNP Q46944
C	-14	HIS	-	EXPRESSION TAG	UNP Q46944

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-13	HIS	-	EXPRESSION TAG	UNP Q46944
C	-12	HIS	-	EXPRESSION TAG	UNP Q46944
C	-11	HIS	-	EXPRESSION TAG	UNP Q46944
C	-10	SER	-	EXPRESSION TAG	UNP Q46944
C	-9	SER	-	EXPRESSION TAG	UNP Q46944
C	-8	GLY	-	EXPRESSION TAG	UNP Q46944
C	-7	LEU	-	EXPRESSION TAG	UNP Q46944
C	-6	VAL	-	EXPRESSION TAG	UNP Q46944
C	-5	PRO	-	EXPRESSION TAG	UNP Q46944
C	-4	ARG	-	EXPRESSION TAG	UNP Q46944
C	-3	GLY	-	EXPRESSION TAG	UNP Q46944
C	-2	SER	-	EXPRESSION TAG	UNP Q46944
C	-1	SER	-	EXPRESSION TAG	UNP Q46944
C	0	MET	-	EXPRESSION TAG	UNP Q46944
C	1	GLY	-	EXPRESSION TAG	UNP Q46944
C	69	LYS	LEU	ENGINEERED MUTATION	UNP Q46944
C	142	GLN	GLU	ENGINEERED MUTATION	UNP Q46944
D	-20	MET	-	EXPRESSION TAG	UNP Q46944
D	-19	ALA	-	EXPRESSION TAG	UNP Q46944
D	-18	SER	-	EXPRESSION TAG	UNP Q46944
D	-17	SER	-	EXPRESSION TAG	UNP Q46944
D	-16	HIS	-	EXPRESSION TAG	UNP Q46944
D	-15	HIS	-	EXPRESSION TAG	UNP Q46944
D	-14	HIS	-	EXPRESSION TAG	UNP Q46944
D	-13	HIS	-	EXPRESSION TAG	UNP Q46944
D	-12	HIS	-	EXPRESSION TAG	UNP Q46944
D	-11	HIS	-	EXPRESSION TAG	UNP Q46944
D	-10	SER	-	EXPRESSION TAG	UNP Q46944
D	-9	SER	-	EXPRESSION TAG	UNP Q46944
D	-8	GLY	-	EXPRESSION TAG	UNP Q46944
D	-7	LEU	-	EXPRESSION TAG	UNP Q46944
D	-6	VAL	-	EXPRESSION TAG	UNP Q46944
D	-5	PRO	-	EXPRESSION TAG	UNP Q46944
D	-4	ARG	-	EXPRESSION TAG	UNP Q46944
D	-3	GLY	-	EXPRESSION TAG	UNP Q46944
D	-2	SER	-	EXPRESSION TAG	UNP Q46944
D	-1	SER	-	EXPRESSION TAG	UNP Q46944
D	0	MET	-	EXPRESSION TAG	UNP Q46944
D	1	GLY	-	EXPRESSION TAG	UNP Q46944
D	69	LYS	LEU	ENGINEERED MUTATION	UNP Q46944
D	142	GLN	GLU	ENGINEERED MUTATION	UNP Q46944
E	-20	MET	-	EXPRESSION TAG	UNP Q46944

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-19	ALA	-	EXPRESSION TAG	UNP Q46944
E	-18	SER	-	EXPRESSION TAG	UNP Q46944
E	-17	SER	-	EXPRESSION TAG	UNP Q46944
E	-16	HIS	-	EXPRESSION TAG	UNP Q46944
E	-15	HIS	-	EXPRESSION TAG	UNP Q46944
E	-14	HIS	-	EXPRESSION TAG	UNP Q46944
E	-13	HIS	-	EXPRESSION TAG	UNP Q46944
E	-12	HIS	-	EXPRESSION TAG	UNP Q46944
E	-11	HIS	-	EXPRESSION TAG	UNP Q46944
E	-10	SER	-	EXPRESSION TAG	UNP Q46944
E	-9	SER	-	EXPRESSION TAG	UNP Q46944
E	-8	GLY	-	EXPRESSION TAG	UNP Q46944
E	-7	LEU	-	EXPRESSION TAG	UNP Q46944
E	-6	VAL	-	EXPRESSION TAG	UNP Q46944
E	-5	PRO	-	EXPRESSION TAG	UNP Q46944
E	-4	ARG	-	EXPRESSION TAG	UNP Q46944
E	-3	GLY	-	EXPRESSION TAG	UNP Q46944
E	-2	SER	-	EXPRESSION TAG	UNP Q46944
E	-1	SER	-	EXPRESSION TAG	UNP Q46944
E	0	MET	-	EXPRESSION TAG	UNP Q46944
E	1	GLY	-	EXPRESSION TAG	UNP Q46944
E	69	LYS	LEU	ENGINEERED MUTATION	UNP Q46944
E	142	GLN	GLU	ENGINEERED MUTATION	UNP Q46944
F	-20	MET	-	EXPRESSION TAG	UNP Q46944
F	-19	ALA	-	EXPRESSION TAG	UNP Q46944
F	-18	SER	-	EXPRESSION TAG	UNP Q46944
F	-17	SER	-	EXPRESSION TAG	UNP Q46944
F	-16	HIS	-	EXPRESSION TAG	UNP Q46944
F	-15	HIS	-	EXPRESSION TAG	UNP Q46944
F	-14	HIS	-	EXPRESSION TAG	UNP Q46944
F	-13	HIS	-	EXPRESSION TAG	UNP Q46944
F	-12	HIS	-	EXPRESSION TAG	UNP Q46944
F	-11	HIS	-	EXPRESSION TAG	UNP Q46944
F	-10	SER	-	EXPRESSION TAG	UNP Q46944
F	-9	SER	-	EXPRESSION TAG	UNP Q46944
F	-8	GLY	-	EXPRESSION TAG	UNP Q46944
F	-7	LEU	-	EXPRESSION TAG	UNP Q46944
F	-6	VAL	-	EXPRESSION TAG	UNP Q46944
F	-5	PRO	-	EXPRESSION TAG	UNP Q46944
F	-4	ARG	-	EXPRESSION TAG	UNP Q46944
F	-3	GLY	-	EXPRESSION TAG	UNP Q46944
F	-2	SER	-	EXPRESSION TAG	UNP Q46944

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	-1	SER	-	EXPRESSION TAG	UNP Q46944
F	0	MET	-	EXPRESSION TAG	UNP Q46944
F	1	GLY	-	EXPRESSION TAG	UNP Q46944
F	69	LYS	LEU	ENGINEERED MUTATION	UNP Q46944
F	142	GLN	GLU	ENGINEERED MUTATION	UNP Q46944
G	-20	MET	-	EXPRESSION TAG	UNP Q46944
G	-19	ALA	-	EXPRESSION TAG	UNP Q46944
G	-18	SER	-	EXPRESSION TAG	UNP Q46944
G	-17	SER	-	EXPRESSION TAG	UNP Q46944
G	-16	HIS	-	EXPRESSION TAG	UNP Q46944
G	-15	HIS	-	EXPRESSION TAG	UNP Q46944
G	-14	HIS	-	EXPRESSION TAG	UNP Q46944
G	-13	HIS	-	EXPRESSION TAG	UNP Q46944
G	-12	HIS	-	EXPRESSION TAG	UNP Q46944
G	-11	HIS	-	EXPRESSION TAG	UNP Q46944
G	-10	SER	-	EXPRESSION TAG	UNP Q46944
G	-9	SER	-	EXPRESSION TAG	UNP Q46944
G	-8	GLY	-	EXPRESSION TAG	UNP Q46944
G	-7	LEU	-	EXPRESSION TAG	UNP Q46944
G	-6	VAL	-	EXPRESSION TAG	UNP Q46944
G	-5	PRO	-	EXPRESSION TAG	UNP Q46944
G	-4	ARG	-	EXPRESSION TAG	UNP Q46944
G	-3	GLY	-	EXPRESSION TAG	UNP Q46944
G	-2	SER	-	EXPRESSION TAG	UNP Q46944
G	-1	SER	-	EXPRESSION TAG	UNP Q46944
G	0	MET	-	EXPRESSION TAG	UNP Q46944
G	1	GLY	-	EXPRESSION TAG	UNP Q46944
G	69	LYS	LEU	ENGINEERED MUTATION	UNP Q46944
G	142	GLN	GLU	ENGINEERED MUTATION	UNP Q46944
H	-20	MET	-	EXPRESSION TAG	UNP Q46944
H	-19	ALA	-	EXPRESSION TAG	UNP Q46944
H	-18	SER	-	EXPRESSION TAG	UNP Q46944
H	-17	SER	-	EXPRESSION TAG	UNP Q46944
H	-16	HIS	-	EXPRESSION TAG	UNP Q46944
H	-15	HIS	-	EXPRESSION TAG	UNP Q46944
H	-14	HIS	-	EXPRESSION TAG	UNP Q46944
H	-13	HIS	-	EXPRESSION TAG	UNP Q46944
H	-12	HIS	-	EXPRESSION TAG	UNP Q46944
H	-11	HIS	-	EXPRESSION TAG	UNP Q46944
H	-10	SER	-	EXPRESSION TAG	UNP Q46944
H	-9	SER	-	EXPRESSION TAG	UNP Q46944
H	-8	GLY	-	EXPRESSION TAG	UNP Q46944

*Continued on next page...*



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	-7	LEU	-	EXPRESSION TAG	UNP Q46944
H	-6	VAL	-	EXPRESSION TAG	UNP Q46944
H	-5	PRO	-	EXPRESSION TAG	UNP Q46944
H	-4	ARG	-	EXPRESSION TAG	UNP Q46944
H	-3	GLY	-	EXPRESSION TAG	UNP Q46944
H	-2	SER	-	EXPRESSION TAG	UNP Q46944
H	-1	SER	-	EXPRESSION TAG	UNP Q46944
H	0	MET	-	EXPRESSION TAG	UNP Q46944
H	1	GLY	-	EXPRESSION TAG	UNP Q46944
H	69	LYS	LEU	ENGINEERED MUTATION	UNP Q46944
H	142	GLN	GLU	ENGINEERED MUTATION	UNP Q46944

- Molecule 2 is a DNA chain called DNA (5'-D(P\*CP\*GP\*GP\*GP\*AP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*GP\*CP\*CP\*GP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	22	Total	C	N	O	P	0	0	0
			453	210	90	131	22			
2	K	22	Total	C	N	O	P	0	0	0
			453	210	90	131	22			
2	M	22	Total	C	N	O	P	0	0	0
			453	210	90	131	22			
2	O	22	Total	C	N	O	P	0	0	0
			453	210	90	131	22			

- Molecule 3 is a DNA chain called DNA (5'-D(P\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*TP\*CP\*CP\*CP\*G)-3').

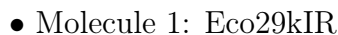
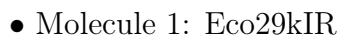
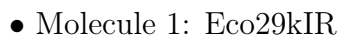
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	22	Total	C	N	O	P	0	0	0
			449	209	85	133	22			
3	L	22	Total	C	N	O	P	0	0	0
			449	209	85	133	22			
3	N	22	Total	C	N	O	P	0	0	0
			449	209	85	133	22			
3	P	22	Total	C	N	O	P	0	0	0
			449	209	85	133	22			

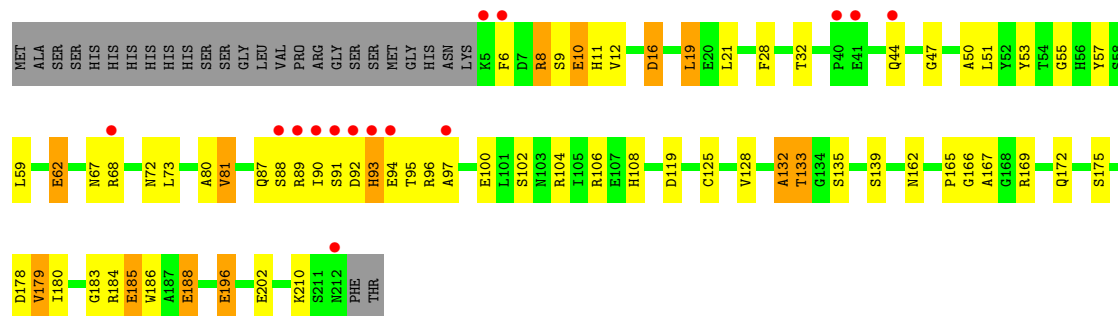
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	74	Total 74	O 74	0	0
4	B	79	Total 79	O 79	0	0
4	C	30	Total 30	O 30	0	0
4	D	28	Total 28	O 28	0	0
4	E	43	Total 43	O 43	0	0
4	F	46	Total 46	O 46	0	0
4	G	66	Total 66	O 66	0	0
4	H	40	Total 40	O 40	0	0
4	I	8	Total 8	O 8	0	0
4	J	8	Total 8	O 8	0	0
4	K	15	Total 15	O 15	0	0
4	L	8	Total 8	O 8	0	0
4	M	8	Total 8	O 8	0	0
4	N	12	Total 12	O 12	0	0
4	O	15	Total 15	O 15	0	0
4	P	13	Total 13	O 13	0	0

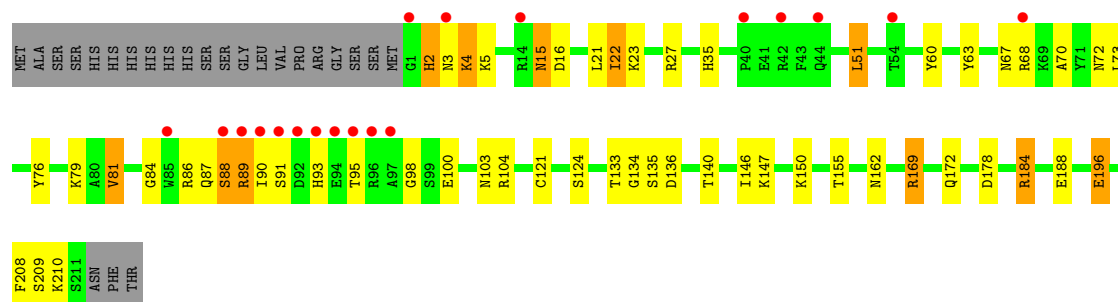


- Molecule 1: Eco29kIR

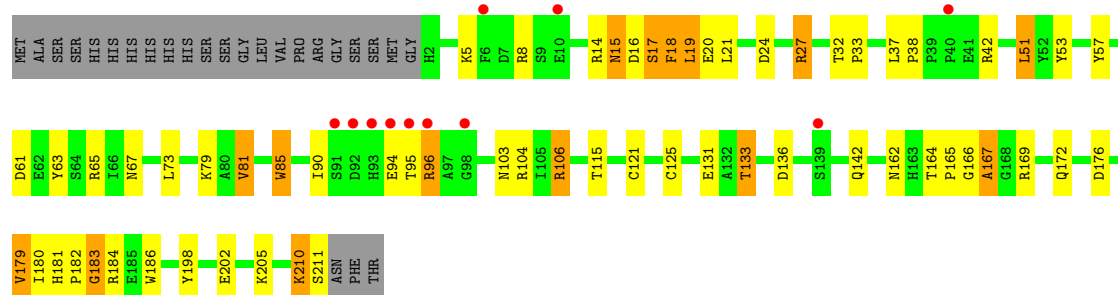




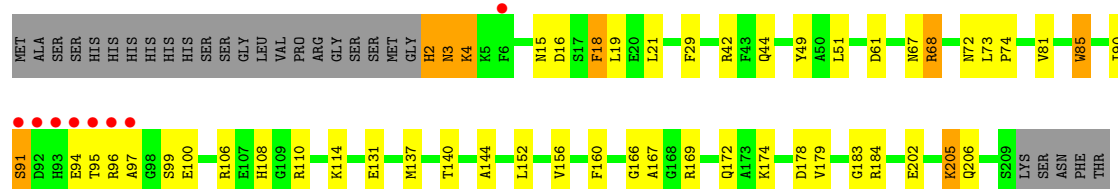
• Molecule 1: Eco29kIR



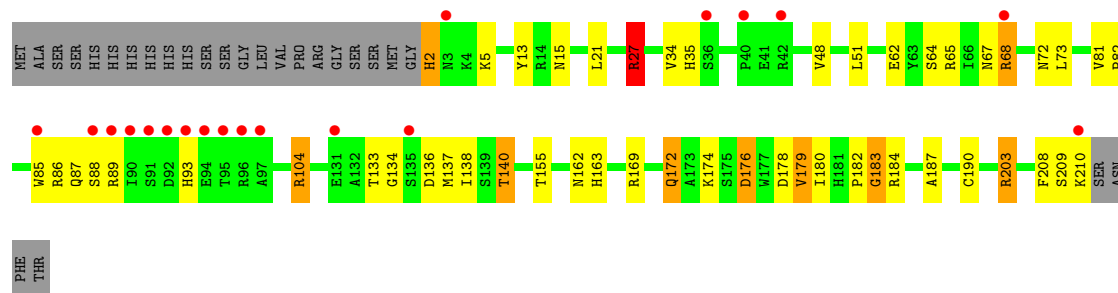
• Molecule 1: Eco29kIR



• Molecule 1: Eco29kIR



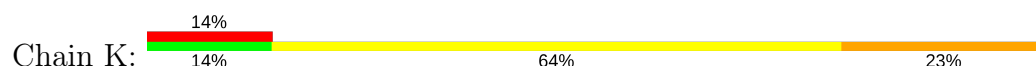
• Molecule 1: Eco29kIR



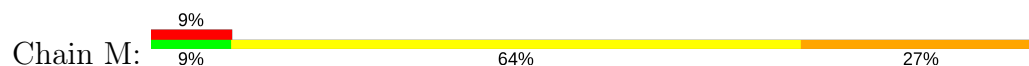
- Molecule 2: DNA (5'-D(P\*CP\*GP\*GP\*GP\*AP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*C  
P\*CP\*GP\*CP\*CP\*GP\*C)-3')



- Molecule 2: DNA (5'-D(P\*CP\*GP\*GP\*GP\*AP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*C  
P\*CP\*GP\*CP\*CP\*GP\*C)-3')



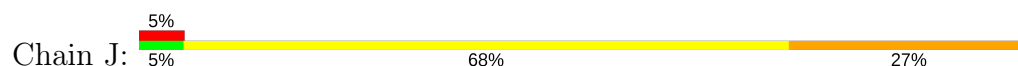
- Molecule 2: DNA (5'-D(P\*CP\*GP\*GP\*GP\*AP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*C  
P\*CP\*GP\*CP\*CP\*GP\*C)-3')

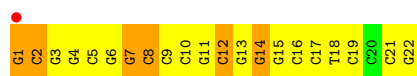


- Molecule 2: DNA (5'-D(P\*CP\*GP\*GP\*GP\*AP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*C  
P\*CP\*GP\*CP\*CP\*GP\*C)-3')



- Molecule 3: DNA (5'-D(P\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*C  
P\*CP\*TP\*CP\*CP\*CP\*G)-3')

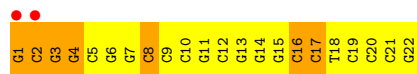




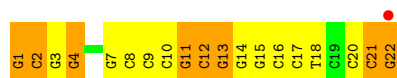
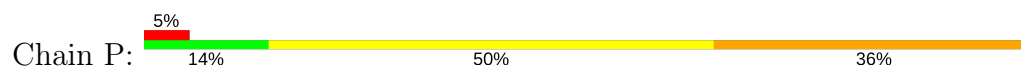
● Molecule 3: DNA (5'-D(P\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*TP\*CP\*CP\*CP\*G)-3')



● Molecule 3: DNA (5'-D(P\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*TP\*CP\*CP\*CP\*G)-3')



● Molecule 3: DNA (5'-D(P\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*TP\*CP\*CP\*CP\*G)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.86Å 101.46Å 144.40Å 90.00° 110.19° 90.00°	Depositor
Resolution (Å)	50.00 – 2.50 44.80 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (50.00-2.50) 99.5 (44.80-2.50)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.28 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, $R_{free}$	0.211 , 0.274 0.219 , 0.275	Depositor DCC
$R_{free}$ test set	4677 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.9	Xtriage
Anisotropy	0.734	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 55.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	17406	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 52.34 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 4.9318e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	1.04	0/1725	0.95	2/2334 (0.1%)
1	B	0.98	0/1738	0.93	2/2357 (0.1%)
1	C	0.88	0/1698	0.89	1/2302 (0.0%)
1	D	0.88	2/1682 (0.1%)	0.87	0/2279
1	E	0.95	0/1713	0.95	3/2325 (0.1%)
1	F	0.97	2/1723 (0.1%)	0.91	1/2333 (0.0%)
1	G	1.01	1/1712 (0.1%)	0.95	1/2322 (0.0%)
1	H	0.93	1/1708 (0.1%)	0.88	1/2317 (0.0%)
2	I	1.57	5/508 (1.0%)	2.28	34/782 (4.3%)
2	K	1.48	6/508 (1.2%)	2.16	26/782 (3.3%)
2	M	1.60	5/508 (1.0%)	2.24	29/782 (3.7%)
2	O	1.53	5/508 (1.0%)	2.34	38/782 (4.9%)
3	J	1.53	3/502 (0.6%)	2.22	30/772 (3.9%)
3	L	1.40	3/502 (0.6%)	2.44	41/772 (5.3%)
3	N	1.52	2/502 (0.4%)	2.38	38/772 (4.9%)
3	P	1.61	6/502 (1.2%)	2.41	35/772 (4.5%)
All	All	1.11	41/17739 (0.2%)	1.40	282/24785 (1.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	F	0	1
1	G	0	2
1	H	0	2
All	All	0	7

The worst 5 of 41 bond length outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	13	DG	C3'-O3'	-8.00	1.33	1.44
3	N	12	DC	N3-C4	7.61	1.39	1.33
2	O	10	DC	N3-C4	7.20	1.39	1.33
3	J	9	DC	C3'-O3'	-6.78	1.35	1.44
3	L	14	DG	N7-C5	-6.68	1.35	1.39

The worst 5 of 282 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	L	13	DG	O4'-C1'-N9	15.04	118.53	108.00
2	K	14	DG	O4'-C1'-N9	-13.64	98.45	108.00
2	I	13	DG	O4'-C1'-N9	13.26	117.28	108.00
2	O	5	DA	O4'-C1'-N9	-13.12	98.82	108.00
2	K	13	DG	O4'-C1'-N9	12.81	116.97	108.00

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	85	TRP	Peptide
1	B	183	GLY	Peptide
1	F	183	GLY	Peptide
1	G	183	GLY	Peptide
1	G	96	ARG	Peptide

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1677	0	1578	49	1
1	B	1687	0	1567	65	0
1	C	1649	0	1540	57	0
1	D	1634	0	1524	56	0
1	E	1663	0	1544	57	0
1	F	1674	0	1578	57	0
1	G	1663	0	1574	41	1
1	H	1658	0	1544	50	0
2	I	453	0	243	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	K	453	0	243	8	0
2	M	453	0	243	12	1
2	O	453	0	243	15	1
3	J	449	0	244	7	0
3	L	449	0	244	11	0
3	N	449	0	244	14	0
3	P	449	0	244	11	0
4	A	74	0	0	4	0
4	B	79	0	0	7	0
4	C	30	0	0	1	0
4	D	28	0	0	3	0
4	E	43	0	0	8	0
4	F	46	0	0	4	0
4	G	66	0	0	0	0
4	H	40	0	0	1	0
4	I	8	0	0	0	0
4	J	8	0	0	0	0
4	K	15	0	0	1	0
4	L	8	0	0	0	0
4	M	8	0	0	0	0
4	N	12	0	0	0	0
4	O	15	0	0	2	0
4	P	13	0	0	4	0
All	All	17406	0	14397	399	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 13.

The worst 5 of 399 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:2:HIS:HB2	1:E:3:ASN:HA	1.18	1.17
1:E:121:CYS:HB2	4:E:408:HOH:O	1.45	1.15
1:C:87:GLN:HB3	1:C:88:SER:CA	1.82	1.09
1:A:140:THR:HG21	1:H:140:THR:HG22	1.39	1.05
1:H:104:ARG:HG2	1:H:104:ARG:HH11	1.21	1.03

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:M:1:DC:OP1	2:O:21:DG:OP2[1_556]	1.97	0.23
1:A:133:THR:OG1	1:G:61:ASP:OD1[2_656]	2.05	0.15

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	211/235 (90%)	194 (92%)	11 (5%)	6 (3%)	6	8
1	B	211/235 (90%)	197 (93%)	8 (4%)	6 (3%)	6	8
1	C	206/235 (88%)	189 (92%)	12 (6%)	5 (2%)	7	11
1	D	206/235 (88%)	188 (91%)	13 (6%)	5 (2%)	7	11
1	E	209/235 (89%)	189 (90%)	13 (6%)	7 (3%)	4	6
1	F	208/235 (88%)	192 (92%)	12 (6%)	4 (2%)	9	15
1	G	206/235 (88%)	191 (93%)	12 (6%)	3 (2%)	12	21
1	H	207/235 (88%)	187 (90%)	17 (8%)	3 (1%)	13	23
All	All	1664/1880 (88%)	1527 (92%)	98 (6%)	39 (2%)	7	11

5 of 39 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	89	ARG
1	A	93	HIS
1	B	4	LYS
1	B	95	THR
1	C	88	SER

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	170/201 (85%)	156 (92%)	14 (8%)	13	25
1	B	170/201 (85%)	157 (92%)	13 (8%)	15	29
1	C	167/201 (83%)	151 (90%)	16 (10%)	10	18
1	D	165/201 (82%)	151 (92%)	14 (8%)	12	23
1	E	164/201 (82%)	149 (91%)	15 (9%)	11	21
1	F	172/201 (86%)	155 (90%)	17 (10%)	9	17
1	G	171/201 (85%)	151 (88%)	20 (12%)	6	12
1	H	168/201 (84%)	151 (90%)	17 (10%)	9	17
All	All	1347/1608 (84%)	1221 (91%)	126 (9%)	10	19

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	210	LYS
1	E	150	LYS
1	H	93	HIS
1	E	15	ASN
1	E	68	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	162	ASN
1	E	44	GLN
1	H	67	ASN
1	E	2	HIS
1	E	103	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	213/235 (90%)	0.54	14 (6%) 19 19	27, 42, 96, 111	0
1	B	213/235 (90%)	0.47	14 (6%) 19 19	28, 42, 92, 109	0
1	C	208/235 (88%)	0.57	19 (9%) 10 10	35, 56, 92, 118	0
1	D	208/235 (88%)	0.52	15 (7%) 16 16	34, 54, 91, 111	0
1	E	211/235 (89%)	0.57	19 (9%) 10 10	29, 46, 97, 130	0
1	F	210/235 (89%)	0.36	11 (5%) 28 29	30, 48, 86, 107	0
1	G	208/235 (88%)	0.37	8 (3%) 41 43	27, 44, 85, 108	0
1	H	209/235 (88%)	0.71	19 (9%) 10 10	33, 52, 98, 133	0
2	I	22/22 (100%)	0.60	2 (9%) 10 10	31, 74, 142, 183	0
2	K	22/22 (100%)	0.68	3 (13%) 3 3	30, 73, 133, 155	0
2	M	22/22 (100%)	0.42	2 (9%) 10 10	32, 65, 124, 142	0
2	O	22/22 (100%)	0.38	1 (4%) 34 36	29, 66, 118, 126	0
3	J	22/22 (100%)	0.47	1 (4%) 34 36	33, 77, 129, 149	0
3	L	22/22 (100%)	0.52	1 (4%) 34 36	30, 80, 144, 156	0
3	N	22/22 (100%)	0.42	2 (9%) 10 10	28, 72, 130, 172	0
3	P	22/22 (100%)	0.59	1 (4%) 34 36	27, 71, 119, 132	0
All	All	1856/2056 (90%)	0.52	132 (7%) 17 17	27, 49, 104, 183	0

The worst 5 of 132 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	90	ILE	16.7
1	A	91	SER	13.0
1	G	93	HIS	11.7
1	E	91	SER	11.2
1	H	91	SER	10.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.